

# code of federal regulations

---

**Patents,  
Trademarks, and  
Copyrights**

**37**

---

**Revised as of July 1, 1994**

**CONTAINING  
A CODIFICATION OF DOCUMENTS  
OF GENERAL APPLICABILITY  
AND FUTURE EFFECT**

**AS OF JULY 1, 1994**

*With Ancillaries*

Published by  
the Office of the Federal Register  
National Archives and Records  
Administration

as a Special Edition of  
the Federal Register



# § 1.822

to satisfy one or more of the requirements of 35 U.S.C. 112. Further, the grant of a patent on an application that is subject to the requirements of §§ 1.821 through 1.825 shall constitute a conclusive presumption that said patent complies with the requirements of §§ 1.821 through 1.825.

(j) Envelopes containing only application papers, computer readable forms and fees filed under this section should be marked "Box SEQUENCE."

[55 FR 18245, May 1, 1990, as amended at 58 FR 4348, Jan. 14, 1993]

## § 1.822 Symbols and format to be used for nucleotide and/or amino acid sequence data.

(a) The symbols and format to be used for nucleotide and/or amino acid sequence data shall conform to the requirements of paragraphs (b) through (p) of this section.

(b) The code for representing the nucleotide and/or amino acid sequence characters shall conform to the code set forth in the tables in paragraphs (b)(1) and (b)(2) of this section. No code other than that specified in this section shall be used in nucleotide and amino acid sequences. A modified base or amino acid may be presented in a given sequence as the corresponding unmodified base or amino acid if the modified base or amino acid is one of those listed in paragraphs (p)(1) or (p)(2) of this section and the modification is also set forth elsewhere in the Sequence Listing (for example, FEATURES § 1.823(b)(2)(ix)). Otherwise, all bases or amino acids not appearing in paragraphs (b)(1) or (b)(2) of this section shall be listed in a given sequence as "N" or "Xaa," respectively, with further information, as appropriate, given elsewhere in the Sequence Listing.

### (1) Base codes:

Symbol	Meaning
A .....	A; adenine
C .....	C; cytosine
G .....	G; guanine
T .....	T; thymine
U .....	U; uracil
M .....	A or C
R .....	A or G
W .....	A or T/U
S .....	C or G
Y .....	C or T/U
K .....	G or T/U

# 37 CFR Ch. I (7-1-94 Edition)

Symbol	Meaning
V .....	A or C or G; not T/U
H .....	A or C or T/U; not G
D .....	A or G or T/U; not C
B .....	C or G or T/U; not A
N .....	(A or C or G or T/U) or (unknown or other)

## (2) Amino acid three-letter abbreviations:

Abbreviation	Amino acid name
Ala .....	Alanine
Arg .....	Arginine
Asn .....	Asparagine
Asp .....	Aspartic Acid
Asx .....	Aspartic Acid or Asparagine
Cys .....	Cysteine
Glu .....	Glutamic Acid
Gln .....	Glutamine
Glx .....	Glutamine or Glutamic Acid
Gly .....	Glycine
His .....	Histidine
Ile .....	Isoleucine
Leu .....	Leucine
Lys .....	Lysine
Met .....	Methionine
Phe .....	Phenylalanine
Pro .....	Proline
Ser .....	Serine
Thr .....	Threonine
Trp .....	Tryptophan
Tyr .....	Tyrosine
Val .....	Valine
Xaa .....	Unknown or other

(c) A nucleotide sequence shall be listed using the one-letter code for the nucleotide bases, as in paragraph (b)(1) of this section.

(d) The amino acids corresponding to the codons in the coding parts of a nucleotide sequence shall be typed immediately below the corresponding codons. Where a codon spans an intron, the amino acid symbol shall be typed below the portion of the codon containing two nucleotides.

(e) The amino acids in a protein or peptide sequence shall be listed using the three-letter abbreviation with the first letter as an upper case character, as in paragraph (b)(2) of this section.

(f) The bases in a nucleotide sequence (including introns) shall be listed in groups of 10 bases except in the coding parts of a sequence. Leftover bases, fewer than 10 in number, at the end of noncoding parts of a sequence shall be grouped together and separated from adjacent groups of 10 or 3 bases by a space.

(g) The bases in the coding parts of a nucleotide sequence shall be listed as triplets (codons).